The Relative Importance of the Vadose Zone in Multimedia Risk Assessment Modeling Applied at a National Scale: An Analysis of Benzene Using 3MRA

Justin E. Babendreier, Office of Research & Development, National Exposure Research Laboratory – Athens, GA babendreier.justin@epa.gov

U.S. Environmental Protection Agency

US EPA Office of Research and Development

US EPA ARCHIVE DOCUMENT

Research Objectives & Strategy

- **Develop Generic Hardware and Software Tools for Conducting Model Evaluation Research & Model Simulation Tasking for Windows & Linux Based Systems**
  - Design & Construct PC-Based Supercomputer (i.e. SuperMUSE) (i.e. Hardware)
  - Create Distributed Parallel Computing Program Management Toolset (i.e. Software)
- **Develop/Expand Underlying Science & Tools for Model Evaluation Tasks**
  - Model Uncertainty & Sensitivity Analysis Methods & Guidance
  - Tool Integration into Framework for Environmental Modeling
    - Facilitate Multiple-Model Integration, Model Comparisons
    - Facilitate UA/SA Model Evaluation Tasks, Data Interpretation
- **Application Migration Strategy**
  - Become a Key Component of EPA's Future Modeling Frameworks (e.g. FRAMES)

...the ultimate medium for convergence of core science, applied science, technology, and regulation.

Problem Background

Elucidating uncertainty and sensitivity structures in environmental models can be a difficult task, even for low-order, single-medium constructs driven by a unique set of site-specific data.

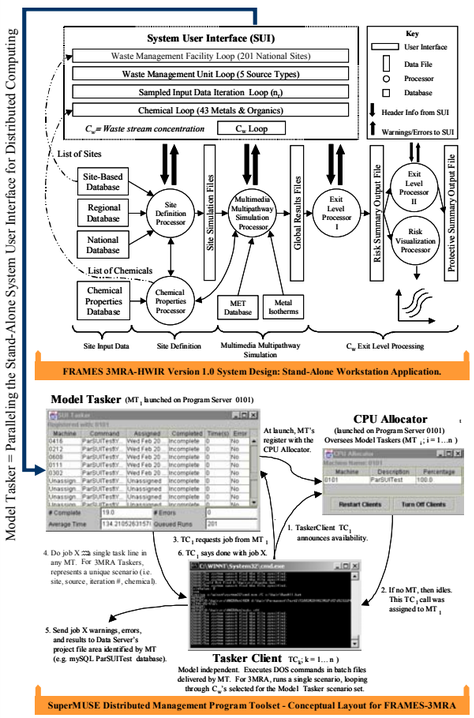
Quantitative assessment of integrated, multimedia models that simulate hundreds of sites, spanning multiple geographical and ecological regions, will ultimately require a comparative model evaluation approach using several techniques, coupled with sufficient computational power.

The Framework for Risk Analysis in Multimedia Environmental Systems - Multimedia, Multipathway, and Multireceptor Risk Assessment (FRAMES-3MRA) is an important software code being developed by EPA for use in assessment of hazardous waste management facilities.

The 3MRA model encompasses over 700 variables, 185 of which are stochastic (i.e. statistical distributions). A characteristic of uncertainty and sensitivity analyses (UA/SA) for very high order models (VHOMs) like 3MRA is their need for significant computational capacity to perform relatively redundant simulations (e.g. need millions upon millions of similar simulations to assess).

While UA/SA is emerging as a critical area for environmental model evaluation, resources for Windows-based, PC-based modeling have often been limited by an associated lack of supercomputing capacity. Equally, higher-order UA/SA algorithms warrant investigation.

Uncertainty & Sensitivity Analysis Software Tool Development: Paralleling Typical Stand-Alone Model Applications for Distributed Computing



FRAMES 3MRA-HWIR Version 1.0 System Design: Stand-Alone Workstation Application.

SuperMUSE Distributed Management Program Tookset - Conceptual Layout for FRAMES-3MRA

DISTRIBUTED MANAGEMENT PROGRAM SOFTWARE - CONCEPTUAL LAYOUT

The CPU Allocator and Tasker Client are model independent. A Model Tasker is model dependent, and deconstructs a model's system user interface to generate a set of tasks (e.g. individual model simulations) amenable to distributed processing. A Model Tasker was developed for 3MRA, identified in here as the SUITasker. Several Model Taskers can be active. The Java toolset is readily extended to Linux by simply recompiling the 3MRA input/output dll called by the SUITasker.

CPU Allocator

The CPU Allocator accepts job descriptions from one or several Model Taskers, and provides proportional load balancing across active Taskers. It functions as a TCP/IP server that accepts Model Tasker scenario set descriptions, and randomly assigns Tasker Clients to Model Taskers when clients indicate that they are free to execute tasks.

Model Tasker

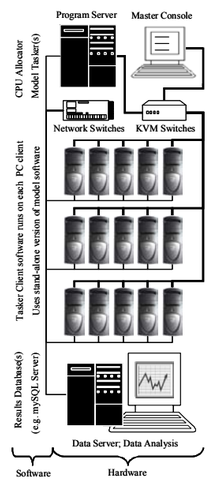
The SUITasker reads a stand-alone 3MRA header file (created in stand-alone mode) to define the overall scenario set to be simulated. It functions as a TCP/IP server that accepts Tasker Clients directed to it by the CPU Allocator. Providing equivalent stand-alone scenario looping, the task list is created, managed, and updated with various statistics to track job performance. It maintains a time-assigned queuing approach and handles errant clients that unexpectedly fail to complete requested tasking (e.g. manages client power failures without client UPS backup).

Tasker Client

The Tasker Client is loaded on each client PC at start-up. Each Tasker Client then periodically calls the CPU Allocator when not tasked. If no Model Tasker is active, it is told to idle. The Tasker Client has no user interface, and functions as a TCP/IP client for the CPU Allocator and active Model Taskers. The client software will connect to the CPU Allocator, receive a Model Tasker machine ID, disconnect, and then connect to the assigned Model Tasker. It then receives a command from the Model Tasker with associated files to execute. In the case of the 3MRA SUITasker, this is a master batch file and a series of 3MRA header files, one for each C_w selected. The file set is first written to the client disk, and the model is then executed in batch sequence.

Providing key connections to the back-end data server, two auxiliary Java applications were also created for client job processing. These also have no user interface. For 3MRA, the tools represent additional calls within the batch file scheme, for each C_w header file executed. The first, a Process Messages Tool, reads normally produced warning and error files and, via JDBC, updates the mySQL server identified by the SUITasker. The second, a Site Summary Tool, extracts results from individual simulation input and output files, storing user-selected variables defined by a delimited script file. Importantly, the auxiliary Site Summary Tool facilitates sensitivity and uncertainty analysis processing for all available input and output data.

Associated Hardware Tool Development: SuperMUSE → *Supercomputer for Model Uncertainty and Sensitivity Evaluation*



SUPERMUSE PARALLEL COMPUTING CLUSTER

A fundamental characteristic of uncertainty and sensitivity analyses is their need for high levels of computational capacity to perform massively redundant model simulations. To facilitate model evaluation tasks for EPA's modelling systems, ORD has recently developed a 125-GHz Supercomputer for Model Uncertainty and Sensitivity Evaluation (SuperMUSE).

Major Hardware Components

Major components of the SuperMUSE include a front-end program server, a back-end data server, and 121 client PCs with a minimum of 256 MB RAM. A variety of Windows operating systems are supported (i.e. Windows 95, 98, NT, 2000). Interconnections were achieved through use of 16-port Raritan KVM (keyboard, video, mouse) switches, and 24-port Linksys (10/100) network switches branching to a master CISCO 3550-24/2 network switch. The system network protocol is based on TCP/IP. System design currently provides for gigE channel (1000 megabits/sec) data flow to and from servers, and also allows single-user KVM remote access in nearby offices.

Various combinations on the cluster design are easily achieved and depend on financial resources (e.g. client speeds, server storage capacities, etc.). Representing a capacity to support 192 clients, the existing SuperMUSE equipment was acquired for $125,000 in early 2001. This excludes servers and 16 older 333 to 450 MHz processors routed to the project. Optimal purchasing based on 5GHz for client PCs will typically identify 3 to 6 month-old CPU technology. Pre-design considerations include available space, and room heating and cooling capacities.

Future Expansion Plans and Anticipated Uses

SuperMUSE will be expanded soon to 192 PCs totalling 270˜GHz. Plans are to accommodate up to 384 PCs, reaching somewhere between 700 to 1000 GHz. Core long-term needs are to facilitate multiple modelling experiments, allowing for simultaneous simulation of multiple models and system use by several modellers.

Summary of Paralleling Capabilities Provided for 3MRA

- Parallel computing software tools represent a critical aspect of exploiting the capabilities of PC-based computing clusters.
- Fairly small, easy to write, and well suited for this application, the Java toolset developed here readily handled machine and job management tasks over the distributed system.
- At its current 125 GHz capacity, SuperMUSE can complete over 2.5 million 3MRA model simulations per month.
- Added runtime was negligible compared to stand-alone, and significantly reduced human capital costs needed to collate data.

Risk Assessment of Alternative National Benzene Disposal Standards

3MRA simulates multimedia (air, water, soil, sediments), fate and transport, multipathway exposure routes (food ingestion, water ingestion, soil ingestion, air inhalation, etc.), multireceptor exposures (resident, gardener, farmer, fisher, ecological habitats and populations; all with various cohort considerations), and resulting risk (human cancer and non-cancer effects, ecological population and community effects).

3MRA is designed to assess which waste streams can safely be released from existing hazardous waste disposal requirements. Hazardous waste with constituent chemical concentrations less than "exit" levels calculated by 3MRA could be reclassified as non-hazardous solid waste.

THE Question

At what waste stream concentration (Cw) will wastes, when placed in a non-hazardous waste management unit over the unit's life, result in:

1. Fewer than A% of the people living within B distance of the facility with a risk/hazard of C or less, and
2. Fewer than D% of the habitats within E distance of the facility with an ecological hazard less than F,

at G% of facilities nationwide?

Benzene risk assessment based on 5 waste concentration levels, 201 randomly selected sites, 419 site-waste management unit combinations, 5 waste management unit types, and 100 Monte Carlo realizations. (209,500 total simulations)

EXIT LEVEL (Cw_{exit}) SCENARIO DESCRIPTION

Human Cancer Risk:

Radial Distance:
500m

Exposure Pathway:
Summation of all Ingestion and Inhalation Pathways

Receptor Type:
All Human Receptors

Cohort:
All Cohorts

Risk Trigger Level: 1.0E-6
Risk Pop. Protection: 99%
Risk Prot. of Sites in U.S.: 95%

Ecological Hazard:

By Ring and Habitat Group

For: <2000m Distance
Habitat Groups:
Terrestrial, Aquatic, Wetland

Eco. Trigger Level: 1.0
Eco. Pop. Protection: 99%
Eco. Prot. of Sites in U.S.: 95%

Surface Impoundments: 137 Sites, Cw_{exit} = 0.13 ppm

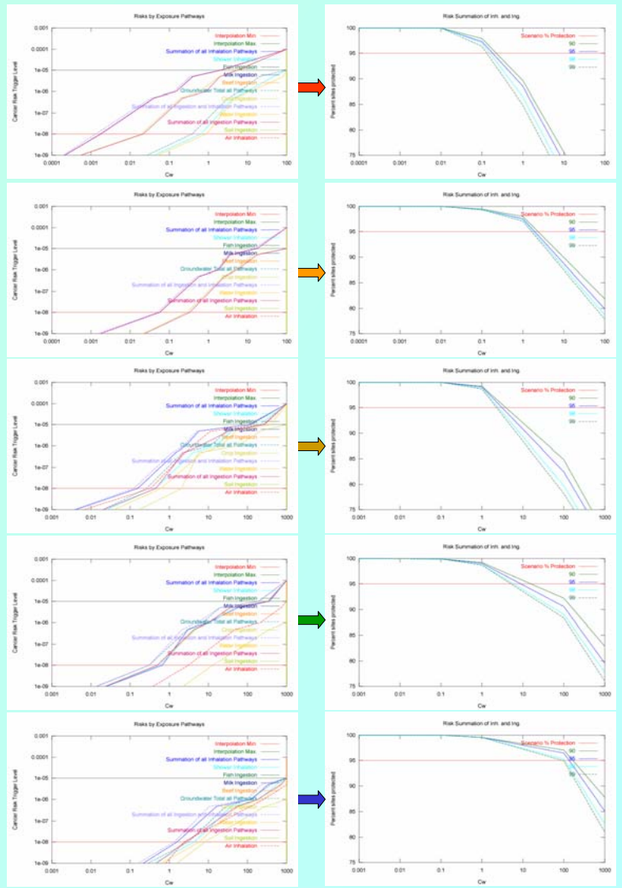
Aerated Tanks: 137 Sites, Cw_{exit} = 1.7 ppm

Land Application Units: 28 Sites, Cw_{exit} = 2.3 ppm

Waste Piles: 61 Sites, Cw_{exit} = 5.3 ppm

Landfills: 56 Sites, Cw_{exit} = 98 ppm

Risk Exposures By Pathway | **Inhalation & Ingestion Summary**



Here we evaluated THE Question *for a specific chemical in terms of:*

How important are the vadose zone and groundwater transport media within the larger multimedia risk assessment problem?

Conclusions of Preliminary Benzene Analysis for 201 National Sites:

- Human Cancer Risk Was The Overall Limiting Concern For All Waste Unit Types. Cw_{exit} ranged from 0.13 to 98 ppm.
- No Concerns Observed for Ecological Hazard for Cw < 100^{+} ppm.
- For Surface Impoundments and Aerated Tanks, Inhalation & Ingestion Via Benzene in Groundwater Were Not Dominant Concern.
- For Land Application Units, Air Inhalation Was As Important As Groundwater-Based Water Ingestion and Shower Inhalation.
- Groundwater Dominated Risks From Waste Piles and Landfills. For Landfills, Air Inhalation Was Important Determinant At Lower Cw.